A SURVEY OF APPLIED PSYCHOLOGICAL SERVICES' MODELS OF THE HUMAN OPERATOR

Arthur I. Siegel and J. Jay Wolf
Applied Psychological Services, Inc.

SUMMARY

A historical perspective is presented in terms of the major features and status of two families of computer simulation models developed by Applied Psychological Services in which the human operator plays the primary role. Both task oriented and message oriented models are included.

Two other recent efforts are summarized which deal with visual information processing. They involve not whole model development but a family of subroutines customized to add the human aspects to existing models.

A global diagram of the generalized model development/validation process is presented and related to 15 criteria for model evaluation.

INTRODUCTION

Our goal in this report is to distill the essence of over 21 years of continuous effort at Applied Psychological Services to develop and validate digital computer models. Our primary area of concentration has been the emulation of systems in which the key element is the human who operates and/or maintains equipment systems. Over 30 projects (contracts) have been devoted to the efforts reported and the models have been applied to systems for all of the military services as well as NASA and industry.

The models are summarized in three families as outlined in Table 1. Three "task oriented models" were developed to simulate man-machine systems. Their major distinguishing feature is the size of the crews which they simulate. They all operate on input data describing a list of tasks or subtasks which the operator(s) or groups of operators perform. Each task is simulated sequentially by the logic of the model implemented by a computer program so as to allow operators to: work independently or together, wait for each other, talk to each other, monitor and operate controls and displays, wait for equipment, skip nonessential subtasks if the operators are busy, make decisions which can alter the subtask sequence, recycle if required in the event of an operator failure, become partially or completely incapacitated, and respond to unexpected failures and emergencies.

The major attribute of the two models in the second family is that they track each message in a communications system. Embedded in each of these, however, is a task oriented module—a miniature of the first model family.

Table 1

Model Families

Family	Family Name	Models
I	Task Oriented Models	One-two Man Model Intermediate Crew Model Large Group Model
II	Message Oriented Models	Tactical Operations Message Handling Field Exercise Monitoring
III	Human Oriented Subroutines for Models	Target Detection & Classification Facility Defense Combat

The paper also describes two recent efforts not directed at development of an entire model--because the model in question had already been developed. In these cases, it was realized after model completion that the human aspects were not adequately simulated. As a result, a group of subroutines was developed to add the realism of a human element in tasks such as scan, detection, and target classification.

All of these models were designed for use in simulating difficult or untried missions--those in which the operator's physical and mental limitations may play an important part in the ability of the man-machine system to perform its function. The conceptual design of each member of each model was based on our unwavering belief that human behavior in a dynamic environment cannot be validly represented by deterministic methods. Last, the paper discusses model validation and a set of validation criteria.

FAMILY OF TASK ORIENTED MODELS

Consider first the family of task oriented models--the mainstay in our arsenal of simulation tools.

All models in this family were prepared to simulate men operating and/or maintaining equipment. All have major simulation variables to reflect the realities of the equipment, the mission itself, and one or more important time functions. Yet they all possess, in addition (and this represents their distinctive feature), psychological and social variables pertaining to the operator or to groups of operators. Examples of these are stress, orientation, proficiency, mental load, and fatigue. Flexibility in simulation is provided in the models through the ability to allow parametric variation of such factors as the speed of the operators, their stress breaking points, and mission time limits. In addition to the more common or system oriented results such as equipment reliability, working hours, and operator failures which one has come to expect from computer models, these three models generate data on personnel performance, morale, cohesiveness, goal orientation, and man-machine system efficiency. All yield computer output tabulations reflective of the man-machine system under study in order to predict system "performance," personnel overloads, periods of unusual stress and excessive delays, distributions of how mission time is spent, a variety of end-of-mission conditions, and implications of manning strategies.

The principal features and differences between the three models are shown in Table 2.

The 1-2 man model (the first entry) simulates one or two operators and accommodates up to 300 individual actions by each operator. Each such operator action which would require a few seconds or minutes of operator performance is simulated by the computer in about 3 milliseconds. Consequently, 100 computer iterations of a maximum task (300 actions for each operator) take about two to four minutes of computer time.

In the intermediate model, a crew of up to 20 men may be simulated. It handles the case of multi-day missions in which the times of individual events

Table 2

Major Features of the Task Oriented Models

NAME	Number of Personnel Simulated	Duration of Mission	Duration of Tasks, Events	Number of Tasks, Events
One-Two Man Model	1-2	minutes, hours	seconds, minutes	up to 300
Intermediate Crew	3-20	a few days	tenths of hours	80 per day
Large Group	20-100	many days	tenths of hours, hours	100 per day

are measured in minutes or hours. This is accomplished by processing tasks performed by groups of one or more men. Here, the computer simulates each of these longer events in about 20 milliseconds. In this case, 100 iterations of a maximum mission (80 crew events per day) for, say, a five day mission, take about 10 to 15 minutes of computer time.

In the largest model, a crew of from 20 to as many as 100 men may be simulated. The mission is composed of work units which may be minutes or hours in duration and the total mission may last for several dozen days. The limit here is principally a practical one based on computer running time.

Table 3 itemizes the principal concepts of the task oriented models. The elements in the table are presented principally to indicate the scope of models. Listed under "Major Human Features and Variables" are the types of functions the model can handle, i.e., the principal variables it considers. The operator-oriented variables selected for incorporation into the models are those which psychologists have determined to be influential on the performance of an individual or of a closed social group. Listed under "Principal Parameters" are the items that the systems analyst who is using the model can vary--that is, he can select values for each parameter for each computer run of the model. Under "Major Outputs" the principal categories of printed computer output are shown.

THE ONE-TWO MAN MODEL

This 1-2 man model has been in active use for almost two decades. It has been tested against real life and against laboratory controlled criteria, and has found to give reasonable, interesting, and valuable results.

The major results from using this model are:

- the probability of success--that is, the percentage of times that the prescribed sequence of subtasks was completed within the time limit.
- the shape of the stress function during the simulation.
- the distribution of time spent working, waiting, and in repeating work not properly performed.

This model is still in active use by a variety of government and commercial users. Several applications have been made by the originators and many by others including landing an aircraft on an aircraft carrier, launching an air-to-air missile, an inflight intercept of an enemy aircraft, and simulation of an inflight refueling operation. These represent both one and two operator simulations.

A preprocessor program has been developed which yields another version of the model. It calculates adjustments to data normally provided as model input prior to simulation. In this case, the effect of radiation exposure on performance time and success probability is determined for man and/or machine degradation.

Table 3

Features of the Task Oriented Models

NAME	1-2 Man Model	Intermediate Crew	Large Group
Major Human Features and Variables	proficiency time stress cohesiveness decision making	proficiency time stress & mental load fatigue & sleep physical capability intelligence learning aspiration supervisors expectation sickness/incapacity	proficiency time stress fatigue norms & goals cohesiveness social pressure learning morale
Principal Equipment/Environment Variables & Features	reliability equipment response	reliability (failures) emergencies communications task postponement consumables hazard level	reliability emergencies communications task postponement consumables mean time to repair
Major Outputs	success probability stress profile work, idle, failure time performance task repetition time tasks failed, ignored	success probability stress profile work, repair, idle, failure time MTBF, MTTR, availability performance adequacy tasks failed, ignored	success probability crew efficiency morale tasks failed, ignored work, idle, failure, repair time
Principal Parameters	operator time limits operator stress threshold operator individuality factor nuclear radiation dose	 operator stress threshold crew work pace work day length acceptable performance level crew qualification level	crew size, increment operator proficiency operator stress threshold acceptable performance level work hours per day crew composition

Table 3 (con't)

Application	aircraft landing inflight intercept sonar inflight refueling missile launching	USCG Patrol Gunboat AN/SQS-76 Sonar	FBM submarine stationary under-water station nuclear missile
Current Use	yes	yes	no

The 1-2 man model was also adapted for subject-to-computer dynamic, on-line interaction. Here, the model simulates the performance of perceptual-motor acts and routine operations while one or two subjects, who are seated at independent graphic video display terminals, perform selected task elements.

THE INTERMEDIATE CREW SIZE MODEL

In the intermediate crew size model, a crew of up to 20 men and multi-day missions may be simulated.

The model, which is heavily group oriented, includes the use of several types of statistical distributions. For example, numbers are drawn from an exponential distribution to determine the time(s) that equipment failures are to be imposed, from a rectangular distribution to determine placement of the emergencies in the list of events each day, from a normal distribution for estimates of mean performance time, and from a poisson distribution to select the number of days duration of sickness of operators. The general simulation sequence is:

- Crew formation--The model identifies each crew member and assigns a value for speed, aspiration, and competence.

- Daily schedule generation--This is done by interspersing prearranged mission events with unforeseen repairs and emergencies.

- Personnel assignment for each event sequentially--Here, the model selects an individual man or a group of men to accomplish the work of each event, ignoring events depending on the essentiality of the events and other factors. The leader of the group is also assigned.

- Event simulation--Calculation of conditions existing prior to the event, and how well and how quickly the assigned men accomplish the work event, and selection of the next course of action.

- Update--Modification of the numerical status of psychosocial and other variables as a result of group performance.

- Output--Selection and printing the value of key variables and summarizing end conditions for each event, each day, each mission iteration, and a summary of all iterations.

The intermediate model was tested to assess its sensitivity and to estimate its validity--that is, the extent to which the model's output agrees with independent criterion data. The mission selected for simulation was that of an 82-foot U.S. Coast Guard boat responsible for patrol of Vietnamese waters. This four-day mission involved a heavy schedule of investigating various river craft, boarding a suspected boat for search operations, navigation, steering, engine monitoring, cleanup, clerical work, preventive maintenance, administrative duties, meal preparation and eating--60 to 65 events in all.

This model has been in almost continuous use since its initial development and validation in 1969. It has been improved by developing a version which simulates equipment, human, and system reliability oriented calculations. As a result, the model yields a number of output numerics believed to possess considerable relevance to human and system availability and reliability prediction. These include: human reliability, availability, and MTTR; equipment reliability, availability, MTTR, and MTBF; and system reliability, availability, and MTTR. This use advances the role of the reliability engineer from that of an actuarial to that of a true system designer or system design advisor who provides an active and ongoing contribution to the total system design and effectiveness assurance process.

This model was also the subject of a set of parametric computer runs for the purpose of developing a set of human tradeoff curves. These were published to show, in handbook format for design engineering use, the relative impact of some human oriented variables on system performance.

THE LARGE GROUP MODEL

In the large group model, the mission is composed of work units, each of which may be minutes or hours in duration, and the total mission may last for several dozen days. Since this model is concerned with group performance, the inputs to the model are principally concerned with group oriented variables salient to behavior. In this mode, variables such as crew morale, cohesiveness, operator orientation, proficiency, performance time, overtime, communications, sickness, and system effectiveness are computed.

In the use of this model, we conceive of supervisors and workers who together form a relatively large crew. In the performance of each job, the computer "selects" the proper number of appropriately skilled men to form a group who "accomplishes" the work in a time and under other conditions which are numerically calculated.

The large group model is the only one of the family in which the computer is programmed to calculate the crew size. This is an optional feature so that simulation runs can be made with the crew composition prespecified, or if left unspecified, simulation will be initiated with what is considered a minimum crew as determined by the logic of the model. Then additional simulations are performed successively--each time with a larger crew. For each increase, the computer selects the most needed man or men to be added to the crew. This process continues uninterrupted until a preset parametric limit on crew size is exceeded, or until the crew reaches a size which eliminates the need for overtime work.

Sensitivity runs made on this model were based on application to a fleet ballistic missile submarine. A series of 10 day missions was simulated with crew sizes which ranged from 33 to 44 men working at five stations, using actual data available during the FBM planning stage. The model was run through cruise operations, stationary submerged operations, and emergency drills which are representative of typical missions. The results compared favorably with actual system mission data. In particular, predictions from the model of system effectiveness, in the composition of the crew, and in its proficiency agreed very well with quantitative data as well as qualitative opinion summarized from interviews with officers of FBM submarines.

Operational validation of this model was completed using data from underseas craft of the 627 class of submarines. Numerous computer simulations of a 21-day mission were made with crew sizes varying from 48 to 61.

It is noted that this model requires extensive data input, and possibly, as a result, this is the only model of the three task oriented models which has had no recent activity.

MESSAGE ORIENTED MODELS

Two models were developed to simulate those aspects of systems whose primary purpose is the operational handling of messages. These models keep track of each message text processed in the system and also simulate the acts and behaviors of operations personnel as they receive, prioritize, code, and enter messages in the system. The models are completely general and allow for the simulation of personnel of different competencies and stress tolerances, along with a variation in message load and content.

These models combine the effects of such features as message generation and queuing, detailed message processing procedure, error rates, and personnel characteristics, along with stochastic variations to yield predictions of system performance. As in the task oriented family, the basic nature of both models is stochastic. As a result, a number of repetitions is required to produce a stable result.

Along with the simulation of human message processing, the models include the simulation of the computer embedded in each of the target systems. Some global information about both of the message processing models is given in Table 4. Both models handle multiple message types of varying priorities.

The first of the models, initiated in 1972, was directed to the simulation of message processing within the Tactical Operations System. TOS is an automated, secure information processing system designed to assist military commanders and their staffs at Field Army, Corps, and Division levels in the conduct of tactical operations.

There are up to four sequences of task elements provided to represent the tasks executed by an operator in performing his duties. Each sequence has the capacity of up to 20 task elements. The model handles up to 6 men of 2 types, 4 types of operator errors, 7 types of messages, 4 message priority classifications, and a shift length of up to 12 hours.

At the start of simulation for a new TOS shift, a backlog subroutine generates data representing messages in the action officer's "in-box" at the start of each iteration. A message generation subroutine develops data representing messages which will arrive during the coming hour. These are merged with the backlog in order by time of arrival, and each message of this hourly message queue is processed in turn by a single selected operator. The operator stress and aspiration conditions applicable to that situation are calculated next. The detailed task element-by-task element simulation for the message and operator selected is accomplished by a subroutine which manipulates mission task analysis data in a way very similar to that used in the 1-2 man model described earlier.

Output from the model includes detailed and summary tabulations including an hourly summary, shift summary, and run summary.

The simulation run summary includes sections for manpower utilization, message processing time, overall efficiency indicator, workload summary, and error summary. In this form, the original sensitivity tests were run, and the model was validated against a set of error data collected from an independent source. A high degree of correspondence with the independent data was found.

In a follow-on effort, the model was modified to operate in an interactive time sharing mode, allowing the experimenter and one or more subjects to interact in a "conversational" mode with the model and to enter data "on line." Various extensions of the original model were also made at this time. A variant of the original model was also included which allowed collection of data during an experiment in which one or more actual operators performed a part of the process and the computer simulated the remainder of the TOS activity.

More recently, the TOS model was adapted for the UNIVAC 1108 computer, and several new capabilities were added which increase the realism of the simulation. It was modified to exchange data with two other independent Army computer models

Table 4

Description of Message Handling Models

System Simulated	Tactical Operational System	Military Exercise Control/ Evaluation System
Program Name	MANMOD	NETMAN
Maximum Number of Men Simulated	6	57
Types of Personnel	2	4
Major Input Parameters	Shift Length Number of Personnel Error Rates Operator Characteristics Speed Precision Aspiration Stress Message Characteristics	Shift Length Number of Personnel Error Rates Operator Characteristics Speed Precision Aspiration Stress Message Characteristics Network Data
Major Output	System Effectiveness Time Worked Operator Stress, Aspiration Message Processing Statistics Errors	System Effectiveness Time Worked Operator Stress, Aspiration Message Processing Statistics

in such a way as to maximize the strong points of each of the models.

The end result is the ability to answer questions such as:

- How does system effectiveness vary as a function of message load, operator level of aspiration, message arrival time distribution, or personnel proficiency?
- What is the effect of increasing or decreasing the manning level or personnel proficiency?
- How much stress was on the operators during the performance of the work of each hour?
- What is the error rate for various message types and for various mannings and personnel attributes within manning?
- How much time was spent, on the average, processing each type of message?

The Army Field Exercise Model

Most of the techniques used in the TOS model were utilized in developing an expanded model for simulating the message handling aspects of Army field exercises in which up to 27 referees, 27 radio operators, and 3 controllers interact in a fixed, closed loop network of communication lines while sharing time on a central computer. Messages introduced into the system are prepared, processed, and entered into the computer by various personnel and delivered to the controllers for evaluation.

Each computer run of the model represents a simulation of up to 10 hours in duration, in which up to 2000 messages can be processed. In this model, each operator type has its own task analysis.

This model has recently been the subject of both sensitivity and validation testing. A series of 59 computer runs enabled statistical test on the effects of a variety of personnel and workload variables, manpower configurations, and task variables. The results were found to be reasonable and appropriate; the most influential variables were operator speed, operator precision, and network configuration. The psychological factors (stress, aspiration level) exerted a much less powerful effect on output.

HUMAN ORIENTED SUBROUTINES

Two other developments have recently been completed which led to the specification of several computer subroutines designed to provide the capability to simulate the personal, psychosocial, and group interactive aspects involved in the target system. The subroutines are designed to be suitable for interfacing with a parent program which simulates other aspects of the system.

The first effort produced four types of different, yet related, computer subroutines or modules. Each of these was conceived to operate as a part of a global computer program whose goal is to simulate the principal ground based, man-machine operations involved in the AN/UPD-X system. In this USAF system, video type displays present replicas in real time from processed data sensed by a side looking radar, mounted in a USAF reconnaissance aircraft.

The subroutines, defined here are those human oriented functions, involved with the capability of simulated operators to perform basic tasks:

- In the SCAN/DETECT Module the operator scans a cathode ray tube (CRT) screen for the presence of targets and detects targets.

- The CLASSIFY Module involves determining which type of target has been detected.

- The DECISION Module simulates operator decision making.

- The COMMUNICATIONS Module involves simulation of inter-operator communications during AN/UPD-X operations.

Each of these subroutines determines the amount of operator time required in the simulated performance of these tasks and determines whether or not the simulated operator(s) performed these tasks adequately (i.e., sucessfully or unsuccessfully).

The AN/UPD-X system was in the design or "evaluation of alternatives" phase during the model development period. As a result, the human oriented subroutines were developed in a sufficiently general way to allow their use during comparative simulation of alternative AN/UPD-X system designs--even those developed by different industrial contractor teams including different AN/UPD-X equipment configurations and diverse operator sequences. Generality was a goal in these module designs--so that the modules will be valid across various equipment and AN/UPD-X system designs developed by several USAF contractors. This objective was achieved in that a user of these modules need only modify inputs to subroutines in order to accommodate system oriented feature differences such as:

- radar coverage area

- CRT display characteristics and size

- target types

- operator ability

- mission time

- communications load

NUCLEAR FACILITY ATTACK SIMULATION

The other effort leading to human effect modules was directed to a model which pitted an attacking force against a force defending a nuclear facility. The hostile intruder attack had been simulated by a hostile attack simulation model which previously had no human behavioral features.

Four features were selected because of their important effect on human performance and were incorporated:

- effects of nuclear radiation
- visual effects of illumination (light level)
- effects of stress
- group cohesiveness effect

MODEL VALIDATION CONCEPTS

Emshoff and Sisson (1970) in a discussion of model validity concluded that: "the only possible evidence of validity for a simulation model that has been developed specifically for a situation is that the model has made satisfactory predictions in the past." They suggested five "preliminary criteria for evaluating first time models" as described by Hermann (1967). These five are identified by an asterisk in the more comprehensive list of 15 criteria for evaluating a simulation model which are displayed in Table 5. These criteria are not necessarily mutually exclusive. Some are overlapping, but all are considered important in some sense and/or for some classes of models.

In order to place these criteria into some perspective and to view the sequential steps through which our models pass, consider Figure 1, which attempts to tie together the various model development/validation phases with these 15 criteria for model evaluation. This figure displays the major steps (large rectangles) from concept and model requirements derivation through the situation in which the model can be considered for decision aiding and eventually for decision making. The 15 numbered vertical arrows, representing the 15 criteria, show that each step in the process yields some measure of utility, feasibility, cost, reasonableness, or validity. It is suggested that a model whose design meets the criteria emanating from the model design box be said to be "suitable" (see lowest oval). A model which is programmed and debugged enters a state here called "testable." After sensitivity testing (and the implementation of corrections to the model as required), the model is said to be "reasonable." Following adequate validation testing, the model is termed "valid" or "useable" for decision aiding and, after the experience of use, the model is "operative," "proven," or "effective." The various types of data and information required as inputs to each phase are shown entering from the left with the resulting documentation outputs exiting to the right.

Table 5

Criteria for Evaluating the Utility of a Computer Model

	Criterion	Definition
1.	Internal consistency	Extent to which the constructs of the model are marked by coherence and similarity of treatment
2.	Indifference to trivial aggregation	Potential of the model to avoid major changes in output when input groupings or conditions undergo insignificant fluctuations
3.	Correct prediction in the extreme (predictive or empirical validity)	Extent of agreement (correctness of predictions) between model and actual performance at very high/low values of conditions
4.	Correct prediction in midrange (predictive or empirical validity)	Like above for middle ranges values of conditions
5.	Construct validity	Theoretic adequacy of the model constructs
6.	Content (variable parameter) validity (Fidelity)*	Extent to which the model's variables/ parameters match real life conditions
7.	Realism or "face validity"*	Extent to which selected content matches each attribute modeled
8.	Richness of output	Number and type of output variables and forms of presentation
9.	Ease of use	Extent to which an analyst can readily prepare data for, apply, and extract understandable results from the model
10.	Cost of development	Value of effort to conceive, develop, test, document, and support
11.	Transportability-generality	Extent to which different systems, missions, and configurations can be simulated
12.	Cost of use	Value of all effort involving use of model including data gathering, input, data processing, and analysis of results

13.	Internal validity*	Extent to which outputs are repeatable when inputs are unchanged
14.	Event or time series validity*	Extent to which simulation predicts event and event patterns
15.	Hypothesis validity*	Extent to which model relationships correspond to similar relationships in the observable universe

* Approaches to validation defined by Hermann (1967)

FOOTNOTE

The task oriented models were originally developed under contract with the Engineering Psychology Programs and Organizational Psychology Programs, Office of Naval Research. Enhancements for radiation (and other decrement effects) and reliability/availability effects were sponsored by the Aeromedical Research Laboratory, Wright-Patterson Air Force Base and the Naval Sea Systems Command, respectively. The message oriented models were sponsored by the U.S. Army Research Institute for the Behavioral and Social Sciences.

The modules relating to hostile attack on nuclear facilities were developed for Sandia Laboratories, and those relating to electronic processed imagery systems were sponsored by Aerospace Medical Research Laboratory and the University of Dayton Research Institute.

References

Emshoff, J.R. & Sisson, R.L. Design and use of computer simulation models. London: MacMillan, 1970.

Hermann, C. Validation problems in games and simulations. Behavioral Science, 1967, 13, 216-230.

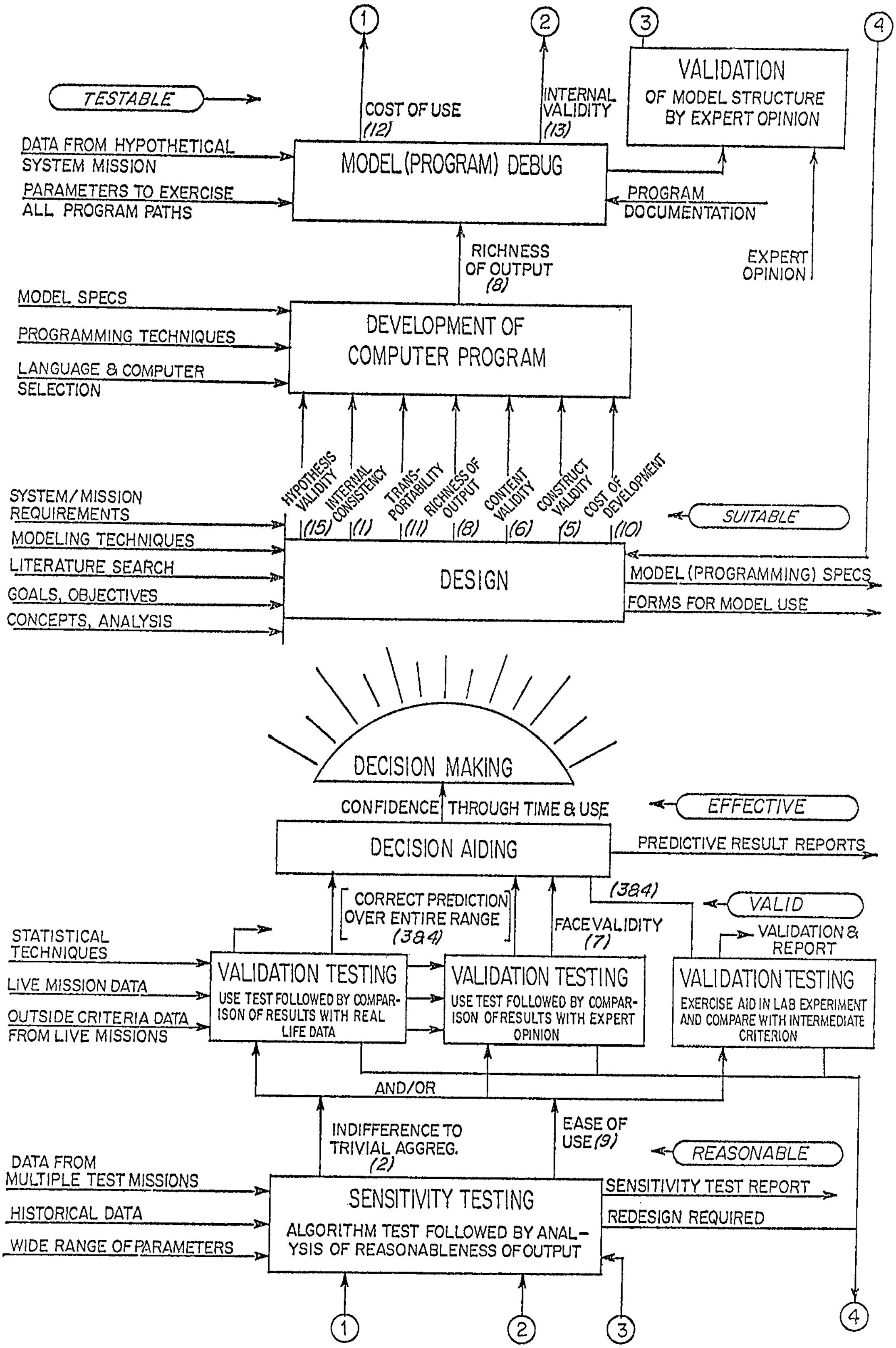

Figure 1.- Steps in model development.